Further insights into the metal ion binding abilities and the metalation pathway of a plant metallothionein from Musa acuminata

Cabral, Augusto C S ; Jakovleska, Jovana ; Deb, Aniruddha ; Penner-Hahn, James E ; Pecoraro, Vincent L ; Freisinger, Eva

Abstract: The superfamily of metallothioneins (MTs) combines a diverse group of metalloproteins, sharing the characteristics of rather low molecular weight and high cysteine content. The latter provides MTs with the capability to coordinate thiophilic metal ions, in particular those with a d10 electron configuration. The sub-family of plant MT3 proteins is only poorly characterized and there is a complete lack of three-dimensional structure information. Building up on our previous results on the Musa acuminata MT3 (musMT3) protein, the focus of the present work is to understand the metal cluster formation process, the role of the single histidine residue present in musMT3, and the metal ion binding affinity. We concentrate our efforts on the coordination of ZnII and CdII ions, using CoII as a spectroscopic probe for ZnII binding. The overall protein fold is analysed with a combination of limited proteolytic digestion, mass spectrometry, and dynamic light scattering. Histidine coordination of metal ions is probed with extended X-ray absorption fine structure spectroscopy and CoII titration experiments. Initial experiments with isothermal titration calorimetry provide insights into the thermodynamics of metal ion binding.

DOI: https://doi.org/10.1007/s00775-017-1513-9

HHS Public Access

Author manuscript

J Biol Inorg Chem. Author manuscript; available in PMC 2018 January 07.

Published in final edited form as:

J Biol Inorg Chem. 2018 January ; 23(1): 91–107. doi:10.1007/s00775-017-1513-9.

Further insights into the metal ion binding abilities and the metalation pathway of a plant metallothionein from *Musa acuminata*

Augusto C. S. Cabral[1], **Jovana Jakovleska**[1], **Aniruddha Deb**[2], **James E. Penner-Hahn**[2], **Vincent L. Pecoraro**[2], and **Eva Freisinger**[1]

[1]Department of Chemistry, University of Zürich, Winterthurerstrasse 190, 8057 Zurich, Switzerland [2]Department of Chemistry, University of Michigan, 930 North University Avenue, Ann Arbor, MI 48109, USA

Abstract

The superfamily of metallothioneins (MTs) combines a diverse group of metalloproteins, sharing the characteristics of rather low molecular weight and high cysteine content. The latter provides MTs with the capability to coordinate thiophilic metal ions, in particular those with a d^{10} electron configuration. The sub-family of plant MT3 proteins is only poorly characterized and there is a complete lack of three-dimensional structure information. Building upon our previous results on the *Musa acuminata* MT3 (musMT3) protein, the focus of the present work is to understand the metal cluster formation process, the role of the single histidine residue present in musMT3, and the metal ion binding affinity. We concentrate our efforts on the coordination of Zn^{II} and Cd^{II} ions, using Co^{II} as a spectroscopic probe for Zn^{II} binding. The overall protein-fold is analysed with a combination of limited proteolytic digestion, mass spectrometry, and dynamic light scattering. Histidine coordination of metal ions is probed with extended X-ray absorption fine structure spectroscopy and Co^{II} titration experiments. Initial experiments with isothermal titration calorimetry provide insights into the thermodynamics of metal ion binding.

Keywords

Metallothionein; Plant; Zinc; Cadmium; Cobalt; Histidine; Dynamic light scattering; EXAFS

Introduction

Metallothioneins (MTs) are rather small proteins (roughly up to 10 kDa) that combine two major characteristics: a high content of the amino acid cysteine and as a consequence a significant binding capacity for soft metal ions, preferentially with a closed *d* electron shell

Correspondence to: Eva Freisinger.

This work is dedicated to Professor Dr. Dr. h.c. Helmut Sigel with the most heartfelt wishes for his 80th birthday and many, many years to come.

Electronic supplementary material: The online version of this article (https://doi.org/10.1007/s00775-017-1513-9) contains supplementary material, which is available to authorized users.

[1–3]. Of biological relevance is the coordination of essential Zn^{II} and Cu^{I} ions, as well as of toxic metal ions such as Cd^{II} and Hg^{II}. The preference of a particular MT form for a certain metal ion is among others determined by the producing organism and a number of intrinsic and extrinsic factors such as the developmental stage, cellular redox status, stress conditions or environmental metal ion exposure. The Cys thiolate groups can function both as terminal and bridging ligands enabling the formation of metal cluster structures in the fully metal ion loaded MT species. There is no unitary function relevant for MTs from all organisms, although metal ion homeostasis, detoxification, and responsiveness to oxidative stress conditions are most frequently reported [4]. MTs are ubiquitous in nature and currently classified into 15 families based on phylogenetic relationships and alignability of amino acid sequences [5]. The most prominent members are the mammalian MTs that form part of the vertebrate MT family (family 1). They have an impressive Cys content of 33% with the thiolate groups serving as sole ligands for seven divalent d^{10} metal ions. The metal ions are arranged in two spatially separated metal-thiolate clusters and coordination of both Zn^{II} or Cd^{II} has been observed in the native protein. Family 15 combines the plant MTs that are further divided into four subfamilies based on the Cys content and the Cys distribution pattern in the primary amino acid sequence. The MT3 proteins represent the plant MT subgroup with the lowest Cys content and hence with the lowest binding capacity for metal ions. As also typical for plant MT1 and MT2 proteins, the MT3 forms feature two Cys-rich regions that are separated by a long Cys-free amino acid linker (~ 30–35 amino acids for the MT3 proteins) and six Cys residues in the C-terminal protein part [6–10]. Distinctive for MT3 proteins is, however, the presence of just four Cys residues in the N-terminal part. Conspicuous is also the presence of one or two histidine residues close to the C-terminal Cys-rich region in roughly 75% of the plant MT3 forms. This is of interest because His can potentially serve as a ligand via its imidazole ring nitrogen, in particular for Zn^{II} ions. So far, coordination of Zn^{II} by His has been structurally proven in two MT forms, one from cyanobacteria and one from wheat [11, 12]. In approximately 1/3 of the His-containing MT3 forms the His is positioned usually four residues before the first C-terminal Cys, while in approximately 2/3 the His resides at the C-terminus. Roughly 20% of these MT3 proteins possess even both His residues [13, 14]. When one or the other His is absent it is nearly exclusively replaced by an asparagine residue (seldom also glutamine or lysine). *Musa acuminata* MT3 (musMT3), the protein investigated in more detail here, features one His in position 46 (H46) and one Asp in the C-terminal position 65 (N65): MSTCGNCDCV DKSQCVKKGN SYGIDIVETE KSYVDEVIVA AEAAE-HDGKC KCGA ACA CTD CKCGN.

Unambiguous identification of Zn^{II} and Cd^{II} coordination by His ligands is considerably more challenging than metal-thiolate binding. While a high resolution solid state structure determined with crystallographic methods can provide a clear “picture” of the coordination site, most other techniques will only yield indications for His coordination. In NMR spectroscopy, [^{15}N,^{1}H]-HSQC spectra optimised for long-range couplings allow one to determine the tautomeric state of the respective His residue, but it needs to be stressed that the observed ^{15}N chemical shift value is not alone a sufficient argument for or against metal coordination [12]. Only the observation of peak splitting due to coupling between ^{15}N and a suitable NMR active nucleus such as ^{113}Cd can provide direct proof of metal binding [15].

Nevertheless, it also needs to be considered in this respect that replacement of Zn^{II} by Cd^{II} will not necessarily produce an isostructural coordination site. Extended X-ray absorption fine structure (EXAFS) spectroscopy is another versatile method frequently applied and delivers information about the number and nature of ligand atoms in the coordination sphere of the respective absorbing metal ion including metal–ligand distances. On the downside, the resulting EXAFS signal is averaged over all absorbing atoms in the sample and hence characterization of different binding environments in a protein or complex with multiple metal ions can be challenging, in particular if the aim is to identify one His ligand next to ten thiolate ligands as in musMT3 [16]. Another approach is the chemical modification of the deprotonated ring nitrogen of His with diethyl pyrocarbonate (DEPC) that can be followed by UV spectroscopy or mass spectrometry [17–19]. The method is widely used for the identification of surface-exposed residues or to tackle the His protonation state [18]. More recently, DEPC was applied to differentiate between metal ion coordinated and free His in MTs [20]. The drawback of DEPC is its rather low specificity, i.e. concomitant modification of the protein N-terminus as well as lysine and tyrosine residues [18]. At increased DEPC-to-protein ratios additional adducts can appear in the mass spectra resulting from modification of both His imidazole nitrogen atoms with DEPC (+ 162 Da) and hydrolysis of the latter (+ 134 Da) [18]. In addition, interpretation of the results can become challenging as incomplete or failed modification can originate from multiple factors. For example, incomplete reactivity occurs when the His residue is buried within the protein or when it is protonated, hydrogen bonded or coordinated to a metal. Generally, all of these factors lead to a non-stoichiometric modification yield.

In MT research, Co^{II} is used as a spectroscopic probe for Zn^{II} that allows the identification of metal-thiolate cluster formation based on shifts of absorption bands in the UV/vis spectra [21, 22]. In addition, the position of the *d–d* transitions in the visible range was used to identify His next to Cys coordination in the zinc finger consensus sequence CP-1 [23]. While the spectrum of the $Co^{II}Cys_4$ binding site is dominated by two bands of nearly equal absorptivity at 695 and 725 nm, exchange of one Cys residue by His leads to a broadening of the absorption envelope that is defined by three bands at 590, 645, and 720 nm (SI). Exchange of a second Cys by His leads to the complete disappearance of the lowest energy transitions resulting in a spectrum with a major band at 640 nm and a less intense band at 570 nm. The same general trend was observed for the DNA binding gene 32 protein (gp32) from bacteriophage T4 that natively contains a Zn^{II} $HisCys_3$ site (SI) [24]. However, the visible part of the Co^{II} spectrum is considerably less well resolved than the corresponding spectrum of the CP-1 Cys-to-His mutant and shows a major, very broadband at 650 nm with unresolved shoulders around 610 and 695 nm. A minor shoulder is discernible at 750 nm. The spectrum of the gp32 His-to-Cys mutant is shifted to lower energy compared to CP-1. The two major bands are positioned at 650 and 690 nm. Again the absorption envelope is quite broad extending, including shoulders, between 610 and 750 nm.

Previously, we presented the successful recombinant production of musMT3 in *Escherichia coli*, its purification and the preliminary characterization of its metal ion binding capabilities [13]. musMT3 can form species with either three or four divalent d^{10} metal ions, the pH stability of metal binding is significantly reduced compared to the mammalian forms, and replacement of Zn^{II} ions from the protein by Cd^{II}, Pb^{II}, and Hg^{II} was shown. In the present

work, we will elaborate on the cluster formation process with a special focus on the observed M_3^{II} and M_4^{II} species, the role of the His residue in metal ion coordination, as well as on metal ion binding affinities using isothermal titration calorimetry (ITC).

Experimental

Chemicals and enzymes

The enzymes and purification systems used for plasmid construction were purchased from Promega AG (Dübendorf, Switzerland), Fermentas (Le Mont-sur-Lausanne, Switzerland), New England Biolabs Inc. (Ipswich, MA, USA), or Roche (Rotkreuz, Switzerland). All buffers and chemicals were ACS grade or comparable from Sigma-Aldrich (Buchs, Switzerland), Merck (Darmstadt, Germany), Chemie Brunschwig (Basel, Switzerland), and Roth AG (Arlesheim, Switzerland). Chelex® 100 resin was from Bio-Rad (Reinach, Switzerland). Plasmid pRK793 for the production of tobacco etch virus (TEV) protease in *E. coli* was a gift from David Waugh (Addgene plasmid # 8827) [25]. All solutions were prepared using Millipore water. If necessary, solutions were saturated with nitrogen or argon. Whenever complete absence of oxygen was required, Millipore water was degassed by three consecutive freeze–thaw cycles under vacuum.

Construction of expression vectors

Construction of the plasmid used for recombinant production of musMT3 in *E. coli* and purification of the protein using the vector pTYB2-musMT3 derived from the IMPACT™-CN system (New England Biolabs Inc.) has been already described [13].

For the N- and C-terminal Cys-rich peptides (C_4 and HC_6, respectively), as well as the three C-terminal H46N N65H (C_6H), N65H (HC_6H), and H46N (C_6) mutants the vector pGEX-4T-1 (GE Healthcare, Uppsala, Sweden) was used. pGEX-4T-1 encodes an N-terminal glutathione *S*-transferase-(GST-) tag as well as a thrombin cleavage site, which will add the two amino acids Gly and Ser to the N-terminus of the target protein. To obtain peptides with their native terminal sequences instead, an additional TEV protease cleavage site was introduced into the vector between the thrombin cleavage site and the target peptide sequences. For this, first, the two plasmids encoding the wild-type N- and C-terminal sequences were constructed using pTYB2-musMT3 as a template and forward primers that include the sequence for the TEV cleavage site (SI) generating pGEX-C_4 and pGEX-HC_6. For the C-terminal mutants C_6 and HC_6H pGEX-HC_6 was used as a template, and the point mutations were created using the primers given in the SI and the method described previously [13]. In the same way, pGEX-HC_6H was used to prepare pGEX-C_6H by point mutation. Identity of the resulting plasmids was verified by DNA sequencing (Microsynth AG, Balgach, Switzerland). Subsequently, the plasmids were transformed into the protease deficient *E. coli* expression strain BL21(DE3).

Protein over-expression and purification

Expression and purification of musMT3 was performed as described [13]. For the expression of the peptides, cells were grown in ampicillin supplemented LB medium (100 μg mL^{-1}) at 37°C and induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at an optical

density of 0.8. After 5 h at 30°C, cells were harvested by centrifugation, re-suspended in phosphate-buffered saline (PBS; 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 140 mM NaCl, 2.7 mM KCl, pH 7.3), and lysed by ultrasonication (Brandson Sonifier 250). The soluble cell extract was applied to a GST column (GSTPrep FF 16/10, GE Healthcare), equilibrated with PBS. After washing, the fusion protein was stripped from the column using 50 mM reduced glutathione and dialysed overnight against 50 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris–HCl), pH 8.0 (SnakeSkin® dialysis tubing, MWCO 3.5 kDa, Thermo Fisher Scientific, Switzerland). Subsequently, the GST-tag was cleaved with TEV protease (peptide to TEV ratio 50:1 w/w) for 1 h at 34°C in the presence of 0.5 mM ethylenediaminetetraacetic acid (EDTA) and 1 mM dithiothreitol (DTT) using slight agitation of the solution (30 rpm). TEV protease was prepared as described before [26]. Separation from the cleaved GST-tag and TEV protease was performed by size exclusion chromatography (SEC) using a HiLoad 16/60 Superdex 30 prep grade column (GE Healthcare) equilibrated with 10 mM ammonium acetate, pH 7.3. The respective peptides were lyophilized and stored at − 80°C for further use.

Protein and metal ion quantification

Protein and peptide concentrations were determined via thiolate group quantification with 2,2′-dithiodipyridine (2-PDS) at pH 4.0 as described [27, 28]. Metal ion concentrations were accessed with flame atomic absorption spectroscopy (F-AAS) on an AA240FS spectrometer (Varian Technologies AG, Zug, Switzerland) in 0.2 M HNO_3 without further digestion of the protein matrix.

Preparation of metal ion saturated Zn^{II}- and Cd^{II}-forms

To obtain fully metal ion saturated musMT3 species, the apo-protein (~ 10 μM, prepared as described below) was mixed with five equivalents of Zn^{II} or Cd^{II} and the pH raised to pH 8.5 with 0.1 M Tris. After 10 min, the excess metal ions were removed by treatment with Chelex® 100 resin (pre-equilibrated in 1 mM Tris–HCl 8.5) for 5 min followed by removal of the resin by centrifugation. Metal ion to protein stoichiometries were confirmed with the 2-PDS assay followed by a F-AAS measurement.

Preparation of apo-musMT3 and apo-peptides

A lyophilized sample of Zn_4musMT3 was dissolved in Arsaturated deionized water and incubated with 50 mM DTT for 30 min at RT. The solution was acidified to pH 2 with 1 M HCl and separated from the released metal ions by SEC [XK 16/20 column (GE Healthcare) self-packed with Sephadex G50 resin (Sigma-Aldrich)] in 10 mM HCl. The eluted apo-musMT3 fraction was saturated with Ar, lyophilized, and kept at − 80°C for further experiments.

For the preparation of the apo-peptides a further optimised protocol was used: the lyophilized peptides were re-suspended in 10 mM HCl, incubated for 10 min at RT with 30 mM tris(2-carboxyethyl)phosphine (TCEP) and applied to a Superdex Peptide 10/300 GL SEC column (GE Health-Care) equilibrated with 10 mM HCl under anaerobic conditions (anaerobic chamber). A single symmetrical peak with increased absorption at 220 nm was collected and immediately used for further experiments.

Metal ion titrations

Metal ion solutions were prepared from $ZnCl_2$, $CdCl_2$, and $CoCl_2 \cdot 6H_2O$ using vacuum degassed and nitrogen saturated deionized water. The Co^{II} solution was additionally rendered oxygen-free by three freeze–thaw cycles using the Schlenk technique. Precise concentrations of metal ion stock solutions were determined by F-AAS. All solutions containing apo-musMT3, apo-peptides or Co^{II} were handled under strictly anaerobic conditions in a glove box equipped with a palladium catalyst and filled with a 5% hydrogen/95% nitrogen gas mix (Coy Lab, Michigan, USA).

For the titration experiments with apo-musMT3, the lyophilized protein was re-dissolved in a buffer containing 1 mM Tris–HCl (pH 7.5) and 10 mM NaCl to give a final concentration of 10 μM for Zn^{II} and Cd^{II} titrations, and 30 μM for the experiments with Co^{II}. For the UV/vis and circular dichroism (CD) measurements with Zn^{II} and Cd^{II}, apo-musMT3 was transferred into a septum sealed cuvette with a path length of 1 cm inside the anaerobic chamber, and metal ion equivalents were added subsequently using a 25 μL syringe with cemented needle (Hamilton Bonaduz AG, Switzerland). After addition of five equiv. of Cd^{II} the CD sample was incubated for 2 min with a small amount (~ 50 μL) of Chelex® 100 resin (Bio-Rad), pre-equilibrated in sample buffer. After centrifugation, another CD spectrum was recorded. For the Co^{II} titrations, one individual sample was prepared for each titration point, transferred into a 100 μL cuvette (path length 1 cm), sealed within the anaerobic chamber and brought to the UV/vis spectrophotometer for data collection.

For the titration of the apo-peptides with Co^{II}, freshly prepared apo-peptides were diluted into a buffer containing 10 mM Tris–HCl (pH 7.8) and 10 mM NaCl to give a peptide concentration of 50 μM, mixed with the required amount of Co^{II} and handled as described above for musMT3.

UV/vis and CD spectroscopy

UV/vis absorption measurements involving musMT3 were carried out on a Cary 500 (Varian AG) scan spectrophotometer, for the peptide titrations with Co^{II} a Cary 60 (Agilent) was used. Scanning speeds were 300 nm min^{-1} and spectral ranges covered 200–500 nm (musMT3 with Zn^{II}, Cd^{II}), 200–800 nm (musMT3 with Co^{II}), as well as 200–900 nm (peptides with Co^{II}). CD spectra were acquired on a J-715 spectropolarimeter (JASCO, Japan) over a spectral range of 200–400 nm with a scan speed of 10 nm min^{-1} and accumulation of 3–4 spectra for each measurement. All UV/vis and CD spectra were recorded at room temperature.

Mass spectrometry

For the measurement of the holo-forms, samples of Zn_4-and Cd_4musMT3 in 10 mM ammonium acetate pH 7.5 were diluted with 50% MeOH. For the measurement of apo-musMT3, a sample of Zn_4musMT3 was desalted using ZipTip C18 (Eppendorf) and diluted with 50% acetonitrile/0.2% formic acid (pH 2). Samples were injected into a quadrupole time-of-flight (TOF) Synapt G2 spectrometer (Waters, UK) and scans accumulated and processed by the software MassLynx 3.5 (Micromass). *m/z* spectra were deconvoluted using the maximum entropy algorithm (Max-Ent1 in MassLynx 3.5). Electrospray parameters

were capillary 2.6 V, cone 50 V, and source temperature 80°C. For matrix-assisted laser desorption/ionization mass spectrometry (MALDI-TOF), the sample was diluted threefold with 0.1% trifluoroacetic acid (TFA) in water. Sample preparation was performed on a steel target using a saturated α-cyano-4-hydroxycinnamic acid solution in 50% acetonitrile and 0.1% TFA as matrix. Mass spectra were recorded on an Ultraflex TOF/TOF spectrometer (Bruker Daltonic GmbH, Germany). The ions were accelerated at 25 kV with pulsed ion extraction time of 100 ns in the positive ion linear mode.

Limited proteolytic digestion with Tritirachium album proteinase K

Cd_3- and Cd_4musMT3 were incubated for 40 min with a 30-fold excess of *T. album* proteinase K in a buffer containing 50 mM Tris–HCl (pH 8.0) and 10 mM $CaCl_2$. Thereafter, the digestion mixture was applied to a Superdex Peptide 10/300 GL SEC column using 60 mM ammonium acetate (pH 7.0) as elution buffer. The collected fractions were further analysed by MALDI-TOF mass spectrometry.

Dynamic light scattering

For dynamic light scattering (DLS) measurements of Cd_3-and Cd_4musMT3 a DynaPro Titan instrument equipped with a temperature controlled microsampler (Wyatt Technology Corporation, Santa Barbara, CA, USA) was used. The measurements were performed with a laser power of 77% at 20°C. 20 μL of a 100 μM sample were centrifuged at 13′200 rpm for 20 min at 20°C prior to the experiment to precipitate any dust that might hamper DLS measurements. At least ten measurements of each sample were recorded and evaluated with the Dynamics® software (Wyatt Technology Corporation).

X-ray absorption spectroscopy

The X-ray absorption spectroscopy (XAS) measurements were performed at the Stanford Synchrotron Radiation Lightsource (SSRL) beamline 9-3 using a liquid nitrogen cooled Si(220) double crystal monochromator with a Rh-coated Zerodur focusing mirror. Spectra were measured as fluorescence excitation spectra, using a 100-element Ge-monolithic solid-state detector; samples were maintained at ~ 10 K throughout the measurements. EXAFS spectra were measured using 10 eV steps in the pre-edge region (9300–9640 eV; 1 s integration time), 0.3 eV steps in the edge region (9640–9690 eV; 1 s integration time), and 0.05 $Å^{-1}$ steps in the EXAFS region ($k = 13.5$ $Å^{-1}$; k^3 weighted integration times from 1 to 15 s integration time). Data from each of the individual channels were examined to confirm that there were no artifacts before averaging; 88 channels were used and a total of seven scans were averaged. Energy calibration was performed by a simultaneous collection of the absorption spectra for the Zn-metal reference foil, with the first inflection point of the foil defined as 9659 eV. Total exposure time for each sample was 3–4 h; comparison of the first and last scans showed no evidence of radiation damage.

Isothermal titration calorimetry (ITC)

Zn_4musMT3 was dialyzed against 10 mM 3-(*N*-morpholino) propanesulfonate buffer (MOPS, pH 7.4) and 10 mM NaCl for 48 h inside an anaerobic chamber replacing the Ar-saturated buffer every 12 h. The metal ion stock solution was prepared by dissolving $CdCl_2$

in the dialysis buffer and the precise metal ion concentration was determined by F-AAS. The titration was carried out with a MicroCal VP-ITC microcalorimeter (Northampton, MA) at 20 °C placed in an anaerobic chamber. The stirred cell contained 1.4 mL of the 14 μM protein solution. The heat released at each of the 25 titration steps was calculated by integrating the differential power measured upon each metal ion addition against the time within a 120-s window. The small heat changes observed under saturating conditions (last injections) arise from dilution and mixing effects and were subtracted from the raw data. The data were fitted to a single-site binding model (Eq. 1) using the program Origin (version 5.0, Micro-Cal) [29, 30].

$$Q = n \times \theta \times M_t \times \Delta H \times V_0 \quad (1)$$

with

Q = heat content of the solution contained in V_0,

n = number of binding sites,

V_0 = cell volume,

M_t = bulk concentration of macromolecule in V_0,

θ = fraction of sites occupied by the ligand,

ΔH = molar heat of ligand binding,

F = fraction of total macromolecule having bound ligands.

Results and discussion

Characterization of reconstituted Zn_4- and Cd_4MT3 with ESI–MS, UV, and CD

Expression and purification of full-length musMT3 was performed as described before with the adaptation that 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer was used during the splicing reaction of the intein fusion protein as well as in the subsequent dialysis step to prevent Tris-adduct formation as reported [13, 31]. In contrast to the previous preparation, the ESI–MS spectrum of apo-musMT3 reveals one species of 7696.5 Da (calc. 6796.6 Da, Fig. 1a) corresponding to the native amino acid sequence without the N-terminal Met residue and with two additional vector-derived C-terminal residues (GlyPro). Fully Zn^{II} and Cd^{II}-loaded species were prepared from apo-musMT3 by addition of an excess of the respective metal ion and subsequent removal of unbound metal ions with Chelex® 100 resin [13]. In both cases, a single species with four metal ions is obtained (Fig. 1b, c): Zn_4musMT3 7047.8 Da (calc. 7049.8 Da) and Cd_4musMT3 7238.0 Da (calc. 7238.2 Da). To further validate the metal ion binding abilities of apo-musMT3, the protein was titrated with incremental amounts of a $ZnCl_2$ solution and UV spectra were recorded (SI). A clear increase of the ligand-to-metal charge transfer (LMCT) band at 230 nm resulting from Zn^{II} coordination by thiolate groups up to the addition of four equiv. of Zn^{II} is observed, which is the maximum binding capacity of the protein for Zn^{II}. The same experiment was performed for Cd^{II} with identical results (Fig. 2a), but additionally provides hints for the cluster

formation process. It has been reported that Cd^{II}-thiolate cluster formation leads to a bathochromic shift of the S→Cd^{II} CT band from initially 250 nm up to roughly 260 nm [32]. To visualize the effect of cluster formation more clearly, Cd^{II}-thiolate coordination was monitored at 250 nm and cluster formation at the lower energy side of the shifted band, i.e. at 270 nm, to reduce the contribution of the 250 nm band. While the absorptivity at 250 nm should increase linearly with Cd^{II} addition (or even slightly less for Cd^{II} equivalents close to the saturation point as most thiolate groups may already be coordinated), the steepness of the absorptivity plot from the 270 nm data should increase with increasing Cd^{II} addition due to the (proposed) onset of cluster formation. Figure 2b shows the plots at both wavelengths and, to enhance the effect, also the quotient $\varepsilon_{250}/\varepsilon_{270}$ is plotted. The decrease in ε_{250} and the increase in ε_{270} is clearly visible after two equiv. of Cd^{II} are added and remains constant after saturation of musMT3 is achieved with four equiv. of Cd^{II}. Hence, two conclusions can be drawn from the data: (1) cluster formation takes place after addition of more than two equiv. of Cd^{II} and (2) cluster formation is not cooperative (otherwise ε_{270} would increase steadily beginning with the first equivalent added). Non-cooperativity of Zn^{II} and Cd^{II} binding to MTs was recently confirmed for human MT1a at physiological pH and proposed to be the general metalation pathway for most MTs [33]. The situation can, however, be different for the reverse process, and cooperative metal ion release was observed for selected MTs [34, 35].

Aiming to corroborate the results obtained with UV spectroscopy, the titration of apo-musMT3 with Cd^{II} was also analysed with circular dichroism (CD) spectroscopy focusing on the part of the spectra to the red of the CD bands originating from the chirality of the peptide backbone around 215 nm. The CD bands in the wavelength range of the LMCT transitions observed in the UV spectra (around 250 nm) are proposed to originate from the peptide-induced asymmetry of the Cd^{II} binding sites [32]. Studies with mammalian MTs showed that cluster formation leads to a change of the initially monophasic to a biphasic shape of the CD spectra; however, the absence of this effect does not exclude a metal-thiolate cluster structure as seen, e.g., for other plant MTs [7, 28, 36]. Also, the CD spectra of the sub-stoichiometric and fully metalated Cd^{II}-MT3 species exclusively show monophasic profiles (Fig. 2c). The band becomes broader with subsequent Cd^{II} addition and the ellipticity increases up to the addition of three Cd^{II} ions, followed by a decrease in ellipticity and a shift to lower energy for the fourth Cd^{II} ion. Although this is the maximum binding capacity of musMT3 according to the ESI–MS spectra and the UV-titrations, addition of a fifth equiv. of Cd^{II} leads to a further change of the CD spectrum before it remains constant. This fifth Cd^{II} ion has, however, a reduced binding affinity as it can be removed by short incubation with the metal ion chelator Chelex® 100 restoring the CD spectrum of Cd_4musMT3. The observation that this fifth Cd^{II} ion does not lead to an increase of the LMCT band in the UV spectrum excludes its coordination to thiolate groups. Nevertheless, its significant influence on the CD bands that originate from the asymmetry of the Cd^{II} binding sites, indicates its structural influence on the coordination environment of the other metal ions. Feasible alternative ligands for Cd^{II} are the carboxylates of Glu and Asp as well as water molecules. The observation of non-thiolate coordination of metal ions in MTs is, however, not unusual [16].

Analysis of the protein fold in Cd_3- versus Cd_4MT3

As described previously, musMT3 species with three as well as four divalent metal ions can be isolated [13]. A structural model was proposed, in which the three metal ions are coordinated separately by the two Cys-rich regions of the protein and addition of a fourth metal ion triggers the formation of a hairpin-like fold, in which the two Cys-rich regions are joined to a single metal cluster. To analyse the feasibility of this model and to visualize the proposed structures, predictions were performed with the program I-TASSER [37]. To obtain a starting model, the I-TASSER algorithm is aimed to find a template or template fragments in the PDB database [38]. With this basis, the full-length model is assembled applying Monte Carlo simulations combined with ab initio modelling of those regions lacking a suitable PDB template match. Calculations are performed, among others, to determine low free-energy states and to remove steric clashes, and after optimization of the hydrogen-bonding network final structures are proposed. No ligand–metal ion interactions are considered. Details of the starting models used and the final models are summarized in the Supplementary Information. The predicted structure of full-length musMT3 is comparable to the model proposed earlier for $\mathrm{M}_4^{\mathrm{II}}\mathrm{musMT3}$ musMT3 in two aspects (Fig. 3a): firstly, the overall protein fold is that of a hairpin and the two Cys-rich regions are so close in space (there is even a Cys7-Cys50 disulfide bridge) that a single metal cluster can be easily envisioned. Secondly, part of the linker region forms a β-sheet involving 17 amino acids and hence 26% of the total amino acids. Previously a β-sheet content of 30–40% was predicted with IR spectroscopy [9]. However, there is also a short stretch of 5 amino acids (8%) in the I-TASSER model that is predicted to form an α-helix, while this was not apparent from IR spectroscopy. To obtain a structural model of the $\mathrm{M}_3^{\mathrm{II}}\mathrm{musMT3}$ species with the proposed independent N- and C-terminal binding sites, the N-terminal sequence including the residues of the linker (Met1-Ala44) as well as the C-terminal sequence with the linker (Val16-Asn65) were analysed separately with I-TASSER. The resulting anti-parallel β-sheet structure of the linker region for both fragments was used to visually overlay the two structures and generate the desired model (Fig. 3b). Again the Cys residues of each region are close together and two separate metal ion binding sites can be easily envisioned. The structure of $\mathrm{M}_4^{\mathrm{II}}\mathrm{musMT3}$ is more globular and the largest distance between opposite side chain atoms amounts to approximately 4.1 nm, while the structure of $\mathrm{M}_3^{\mathrm{II}}\mathrm{musMT3}$ is clearly elongated/rod like with the longest dimension of approximately 3.1 nm.

To analyse the proposed arrangement experimentally, two approaches were used, i.e. dynamic light scattering (DLS) and limited proteolytic digestion. The DLS measurements were performed with the Cd^{II}-bound species as the resulting cluster structures should be more stable according to thermodynamic considerations. Both samples, Cd_3musMT3 and Cd_4musMT3, revealed low polydispersity (13 ± 1 and 11 ± 2%, respectively), which is an indication for a narrow species distribution and hence a uniform sample. The hydrodynamic radii obtained are 2.01 ± 0.18 and 1.45 ± 0.10 nm, respectively, which is in good agreement to the diameter measurements performed with the predicted structures and corroborates the hypothesis made previously.

Limited proteolytic digestion of MTs with proteinase K in combination with SEC and ESI–MS or MALDI-TOF was used previously to investigate the number of metal clusters formed and to identify the participating regions of the primary amino acid sequence [28, 36, 39]. Proteinase K from *Engyodontium album* (previously *Tritirachium album*) is a rather unspecific serine protease that cleaves mainly after hydrophobic and aromatic amino acids. However, it was shown that the protein backbone is to a certain extent protected by the metal-thiolate clusters and as a result, MTs are mainly cleaved in the Cys-free linker regions allowing the identification of domains. In addition, although linker containing domains can experience internal cleavage, the metal-thiolate clusters are usually preserved keeping the resulting peptide fragments tightly enough together to enable the separation of domains by SEC. In this way, it was shown that the metal ions in the Zn_6E_c-1 MT from wheat germ are arranged in two separate metal binding domains, a smaller one formed by the N-terminal Cys-rich region and a larger one by the combined central and C-terminal Cys-rich regions of the protein [36]. Owing to their significantly differing sizes (2.4 versus 4.5 kDa), separation of the digestion mixture with SEC reveals two distinct peaks. MS data show the intact N-terminal Cys-rich region (2.4 kDa) in one as well as the two peptide fragments of the central (1.9 kDa) and C-terminal Cys-rich regions (2.6 kDa) in the second peak fraction.

To evaluate the number of metal ion binding domains, Cd_4- and Cd_3musMT3 were analysed as described above. First, the elution volume in SEC (Superdex Peptide 10/300 GL) for both undigested species was determined as a control. In agreement with the results from DLS, the elution volume of Cd_3musMT3 is smaller than for Cd_4musMT3 (9.9 versus 10.3 mL) reflecting its larger hydrodynamic radius (Fig. 4). SEC of digested Cd_4musMT3 shows a major elution peak at 11.8 mL and a smaller peak at 13.6 mL that cannot be resolved cleanly. Analysis of the peak fraction at 11.8 mL with MALDI-TOF shows the presence of fragments that can be assigned to the N-terminal as well as the C-terminal Cys-rich regions (the latter including His46) each with part of the linker region of various length (Fig. 5). The fragment of 2274.31 Da is the only one that fits the mass of a fragment solely containing residues of the linker region (Lys17-Glu36); however, the assignment is ambiguous as two more fragments have an identical molecular mass (Cys7-Val27, Glu42-Gly64). Digestion of Cd_3musMT3 results in a SEC elution profile that is shifted to smaller peptide size with two overlapping peaks at 12.7 and again 13.6 mL (Fig. 6). MS data of the combined peak fractions show numerous lower molecular mass fragments that can be assigned to all different parts of the protein. Comparing the results for both species, it is intriguing to conclude that the peak at 11.8 mL observed after digestion of Cd_4musMT3 originates from a joined cluster structure connecting the N- and C-terminal parts of the protein. The peak at 12.7 mL resulting from the digestion of Cd_3musMT3 apparently also contains the N- and C-terminal parts of the protein, but as elution occurs in the smaller size fraction the two parts were likely not connected to each other by metal ions. The peak at 13.6 mL might contain the smallest peptide fragments depicted in Fig. 6b. This interpretation is at least in part supported by the Cd^{II}-to-Cys ratios determined with F-AAS and the 2-PDS assay for the determination of thiolate groups. In the peak at 11.8 mL the ratio is 1:2.6, which is close to the ratio expected for a binding site made up of 4 Cd^{II} ions and 10 Cys residues (calc. 1:2.5). However, the ratio in the peak at 12.7 mL is only 1:2.4, while for 3 Cd^{II} ions coordinated to 9 or 10 Cys residues ratios of 1:3.33 or 1:3.0 would be expected. The Cd^{II} content in the

fraction eluting at 13.6 mL is below the detection limit. The significantly larger size of the peptide fragments observed in the MALDI-TOF spectra of Cd_4musMT3 is another indication for a more compact cluster structure that protects the peptide backbone more efficiently against proteolytic cleavage than the presumably elongated and hence more open arrangement in Cd_3musMT3. Accordingly, only rather small peptide fragments are detected after digestion of Cd_3musMT3.

Co^{II} as a spectroscopic probe: cluster formation and role of the His residue in metal ion coordination

As elaborated in the introduction, the detection of metal ion coordination, in particular of what is known as "spectroscopically silent" closed-shell metal ions such as Zn^{II} and Cd^{II}, can be challenging. Unfortunately, as is the case for many MTs, structural methods such as X-ray crystallography and NMR spectroscopy were not successful, probably due to the inherent flexibility of musMT3. Hence we used Co^{II} as a spectroscopic probe to evaluate cluster formation and His contribution to metal binding. Similar to Cd^{II}, also the thiolate-to-Co^{II} charge transfer bands in the UV region of the spectra experience a bathochromic shift when the coordination mode changes from solely terminal thiolate ligands to a mixed terminal/bridging environment and hence a clustered structure. In contrast to Cd^{II} (and Zn^{II}), coordination of Co^{II} to MTs also gives rise to *d–d* transition bands that are equally influenced by changes in the coordination geometry, i.e. cluster formation [21, 22]. Visualization and interpretation of these bands can be more straightforward as no spectral overlap with transitions from the protein scaffold exists in this wavelength range. Similar to the observed metalation pathways for MTs with Zn^{II} and Cd^{II}, coordination of Co^{II} ions to MTs was reported to follow mainly a non-cooperative mechanism [40, 41]. Figure 7a shows the UV/vis range of the absorption spectra upon titration of apo-musMT3 with increasing equivalents of Co^{II} under strictly anaerobic conditions. The *d–d* transitions are observed in the range of 540–800 nm and consist of three major bands centred around 620, 690, and 740 nm. Such an absorption envelope is typical for Co^{II}-substituted MTs as described in the literature with the respective bands lying in the range of 600–620, 680–700, and 730–750 nm [28, 36, 42]. As for the titrations with Zn^{II} and Cd^{II} the molar absorptivity of musMT3 increases up to four equiv. of Co^{II} and remains constant thereafter (Fig. 7b). The band shift is clearest following the position of the central highest *d–d* band around 690 nm, which starts to experience a bathochromic shift upon addition of the third and fourth Co^{II} ion and then remains constant. This indicates that the first two equivalents of Co^{II} are coordinated in individual binding sites by solely terminal but not bridging thiolate ligands. This is quite feasible considering one metal ion bound to the N-terminal Cys-rich region (4 Cys) and one to the C-terminus (involving 4 of the total 6 Cys). Following our interpretation of two separate binding sites in the M_3^{II}-species, coordination of the third Co^{II} ion takes place at the C-terminus, and according to the band shift, involves at least one bridging thiolate group. Assuming the smallest possible cluster size, that is $M^{II}Cys_5X$ (with X being either Cys, His, or even Asp or H_2O), at least two more ligands are required [43]. We deliberately use caution in defining the sixth ligand X, because addition of the fourth metal equivalent (Zn^{II}, Cd^{II}, or Co^{II}) still increases the absorptivity by approximately 10%. This suggests that after addition of three metal equivalents, only nine Cys are involved in binding at the pH studied, while 10 Cys are used for four metal ions. Coordination of the fourth metal ion is again

accompanied by a bathochromic shift of the central *d*–*d* band indicating that further clustering involving additional bridging thiolate ligands takes place.

As described in the introduction, replacement of Cys by His in tetrahedral Co^{II} coordination sites causes a shift of the highest intensity band in the *d*–*d* region of the spectra to higher energy. Such a shift is clearly not obvious from the Co^{II} spectra of musMT3. A comparison with other Co^{II}-MT spectra from the literature reveals the following. Firstly, the coordination of Co^{II} to individual $CoCys_4$ sites in MTs results in closely similar spectra for all published data including musMT3 with a major band located around 690 nm and two flanking bands around 610 and 740 nm being of roughly equal intensity (SI) [28, 36, 42]. Secondly, while this shape of the *d*–*d* transitions envelope is approximately retained in fully metalated Co_4musMT3, for MTs that form metal clusters involving exclusively thiolate ligands the lowest energy band around 740 nm increases relative to the other two bands resulting in spectra similar to the one of CP-1 [28, 42]. Only a single study is available showing Co^{II} data for a His-containing MT, that is the wheat E_c-1 MT [36]. The six metal ions in E_c-1 are coordinated in a $M^{II}Cys\ His_2$ site and a $M^{II}_2Cys_6$ as well as a $M^{II}_3Cys_9$ cluster. In agreement with the literature data on isolated $CoHisCys_3$ and $CoHis_2Cys_2$ sites, the highest energy band in fully metalated Co_6E_c-1 is increased in intensity relative to the two other bands.

In an attempt to lower the Cys:His ratio and hence to facilitate detection of any spectral contribution from Co^{II}-His coordination, the Cys-rich regions of musMT3 were also analysed separately. As the results presented above strongly indicate independent metal ion coordination by the N- and C-terminal Cys-rich regions in $M^{II}_3musMT3$, this approach should be an appropriate model to study the coordination of the first three metal ions. Hence the two wild-type sequences of the Cys-rich regions (denoted C_4 and HC_6, respectively) were cloned, recombinantly produced, and purified as well as three additional mutants of the C-terminal region to be used for spectra comparisons (Table 1).

The H46N (C_6) mutant is devoid of His residues and accordingly the resulting Co^{II} spectrum will be free of any Co^{II}-His contributions. The H46N N65H (C_6H) was produced as roughly 50% of all plant MT3 proteins contain a C-terminal His (only ~ 25% contain His in the position found in musMT3). Approximately 15% of all MT3 protein contain both His residues, reflected by the N65H mutant (HC_6H). The intention for the production of mutants with additional His residues or His at other positions was to increase the probability to observe any absorption bands from Co^{II}-His coordination and hence to have a positive control for data interpretation using peptides with an amino acid sequence as similar as possible to wild-type musMT3. The UV/vis spectra obtained after addition of one and two equiv. Co^{II} are shown in Fig. 8a. Coordination of the first Co^{II} ion should result in a $CoCys_4$ environment in all cases (in particular for sequences C_4 and C_6), and accordingly, the shape of all spectra in the *d*–*d* range is rather similar and comparable to the respective spectra obtained for other full-length MTs as described above. However, after addition of the second Co^{II} equivalent, the spectra of the C-terminal sequences show significant variations. The spectrum of C_6 shows a relative increase of the lower energy band around 750 nm as seen for the Cys-only Zinc-finger sequence CP-1 and for MTs forming exclusively Co^{II}-thiolate clusters (SI). In contrast, C_6H and HC_6H reveal the development of a band around 650 nm

indicative for His coordination as seen in the two CP-1 His mutants as well as in gp32. The differences become even more apparent upon examination of the difference spectra that are obtained by subtracting the Co_1 spectrum from the respective Co_2 spectrum (Fig. 8b). As for the overall shape of the absorption envelope, not only band intensity but also bandwidth is of importance, Gaussian fitting of the visible spectra was performed (Fig. 8c, d and SI). The spectra obtained after addition of one equiv. of Co^{II} (including the one of Co_1C_4) as well as the spectrum of Co_2C_6 can be fitted with three Gaussian peaks centered around 610, 690, and 740 nm. For the fitting of Co_2C_6H, Co_2HC_6H, and also Co_2HC_6, however, a fourth band centered around 650 nm is required, and this might indicate a contribution from Co^{II}-His coordination to the spectra.

Prior to evaluating the data of the wild-type C-terminal sequence HC_6 in more detail, the spectrum of Co_2HC_6H needs to be considered more closely. The shape of its absorption envelope (Fig. 8a) seems to differ significantly and clearly the intensity is increased. Hence on first glance, it is very intriguing to conclude that in Co_2HC_6H both His must play a role as ligands (forming either two $CoHisCys_3$ sites or one $CoHis_2Cys_2$ plus one $CoCys_4$ site), while in Co_2C_6H only one His is involved (thus forming one $CoHisCys_3$ plus one $CoCys_4$ site). Both proposals can be easily (although only roughly) tested. If Co_2HC_6H forms two $CoHisCys_3$ sites, then dividing the spectrum by 2 should result in a spectrum similar to that of $Co_{(2–1)}$C6H. This is apparently not the case as the intensity of the bands at 690 and 740 nm is much higher relative to the intensity of the band at 650 nm (Fig. 8b). On the other hand, If Co_2HC_6H forms one $CoHis_2Cys_2$ and one $CoCys_4$ site, then the $Co_{(2–1)}HC_6H$ spectrum should differ considerably from the $Co_{(2–1)}$C6H spectrum, which represents a $CoHisCys_3$ site. This is only true with respect to the intensity. The shapes of both spectra are, however, quite similar which becomes more apparent when reducing the intensity of the $Co_{(2–1)}HC_6H$ spectrum by 25% (Fig. 8b). In particular, the part of the visible spectrum in the range of ~ 500–680 nm is nearly a perfect match, while the intensity of the $Co(2–1)C_6H$ spectrum between 680–800 nm is slightly higher. This observation triggered the proposal that in Co_2HC_6H only one of the two His is coordinating, while the composition Co_2C_6H actually represents a mixture of two species, one in which the His residue is coordinating, i.e. $Co_2HisCys_5$, and one with exclusive thiolate coordination of Co^{II} ions, i.e. Co_2Cys_6. A nearly perfect fit of the data for $Co_{(2–1)}C_6H$ was obtained when 75% of the $Co_{(2–1)}HC_6H$ and 25% of the $Co_{(2–1)}C_6$ spectra were summed (Fig. 8b).

Having in such a way “defined” the outer limits of spectra defining His (HC_6H) versus Cys (C_6) coordination, the spectrum of $Co_{(2–1)}HC_6$ was fitted in the same way, giving the best approximation of the experimental results using 30% of the $Co_{(2–1)}HC_6H$ and 70% of the $Co_{(2–1)}C_6$ spectra (Fig. 8b). Hence, also here clearly a mixture of species is present.

Having now a set of reference spectra of the separate Cys-rich regions in hand, attempts were made to fit the Co^{II} spectra of musMT3 in the visible range. The spectra of Co_1C_4 and Co_1HC_6 have roughly the same shape and absorptivity as the spectrum of Co_1musMT3 (Fig. 9a) and accordingly, it can be assumed that the first Co^{II} ion is coordinated in the full-length protein in a $CoCys_4$ site. The Co_2musMT3 spectrum does not reveal large similarities to either of the four spectra of the C-terminal Cys-rich sequences after addition of two equiv. of Co^{II} (Fig. 9a, b, and SI), but overlays to a large extent with the sum of the spectra of Co_1C_4

and Co_1HC_6 (or Co_1C_6). Accordingly, it is unlikely that the first two equiv. of Co^{II} are both coordinated in the C-terminal Cys-rich region of musMT3 but rather that one Co^{II} ion is bound to the N-terminus and one to the C-terminus, most likely both in $CoCys_4$ sites. The spectrum of Co_3musMT3 is best described by the sum of the spectra of Co_1C_4 and Co_2HC_6 (or Co_2C_6), although, as said above, the intensity of the absorption envelope has to be reduced by 10%. While the reason for this is not clear, probable explanations could be a reduced number of ligands (which would not be in line with our interpretations), an influence of the linker on the overall geometry of the metal binding sites and hence on the absorptivity, or a contribution of bridging thiolate groups to the overall absorptivity (increasing the band intensity in the Co_4musMT3 species). Nevertheless, the fit involving Co_2HC_6 results in a slightly increased absorptivity below 660 nm, while the fit involving Co_2C_6 shows an increased absorptivity above 700 nm. Hence again, as performed for the spectra of the separate C-terminal peptides, fitting with different ratios of Co_2HC_6H and Co_2C_6 was attempted, indeed improving the fit when 80% of Co_2C_6 and 20% of Co_2HC_6H are considered (Fig. 9a). This result is in line with (1) a mixture of species as seen for the separate C-terminal peptide and (2) with the coordination of three Co^{II} ions in musMT3 in the separate N- and C-terminal Cys-rich regions.

This finding is also interesting in light of the I-TASSER results described above (Fig. 3). For both arrangements, that is a peptide fold with separated Cys-rich regions as well as a fold with connected regions, H46 is oriented away from the agglomerated Cys residues and involved in a hydrogen bond with the glutamic acid residue E42 of 3.4 and 3.5 Å length, respectively. A His residue at position 65, however, should be very close to a potential metal cluster and can be more easily envisioned to participate in metal ion binding. It is merely hypothetical, but not entirely unfeasible that H46 has a structural role in plant MT3 proteins. That being said it needs to be stressed again that I-TASSER does not model metal–ligand interactions and hence the arrangement of H46 in the native metal-bound form might be quite different.

Extended X-ray absorption fine structure (EXAFS) spectroscopy

Zn-EXAFS may provide precise metrical parameters for the metal ion coordination environment. However, as the EXAFS signal is averaged over all different metal ions in the sample, we did not investigate the full-length protein as the His contribution observed for a putative $Zn_4Cys_{10}His$ *versus* a Zn_4Cys_{10} arrangement would be clearly within the error range of the measurement (i.e. $ZnS_{3.75}N_{0.25}$ *versus* ZnS_4). Hence, only the His-containing C-terminal Cys-rich peptide HC_6, which must discriminate between $ZnS_{3.5}N_{0.5}$ and ZnS_4 coordination, was analysed. Even this stoichiometry is likely to be outside the range that can be directly differentiated by EXAFS; however, one can assess whether a mixed S/N coordination sphere is consistent with the collected spectra. The EXAFS data of Zn_2HC_6 is dominated by a single frequency, giving rise to a single peak in the Fourier transform (Fig. 10). Satisfactory data fitting is achieved with both settings, either considering only a single shell of four sulphur ligand or a mixture of sulphur and lighter, e.g., imidazole ligands. The refined Zn–S distance of 2.33 Å is fully consistent with thiolate ligands in a tetrahedral coordination mode around the Zn^{II} ion. The variation in fit quality as a function of the percent nitrogen ligation (SI) is identical to that seen previously for authentic ZnS_4 sites

[44]. This is consistent with the near absence of outer-shell scattering although if there were two Zn^{II} ions per peptide, one with ZnS_4 and one with ZnS_3N ligation, the outer-shell scattering would be expected to have an amplitude similar to that seen in the peptide. The Zn–N distance of 1.99 Å that results from this alternative fit is fully consistent with a tetrahedral Zn^{II} ion in the presence of other sulphur ligands. In summary, while the EXAFS data alone neither provide prove of nor exclude His as a ligand (or any other low-Z ligand such as oxygen) in Zn_2HC_6, they are nevertheless consistent with the structure proposed based on the UV/vis spectra.

Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) allows the quantitative study of various binding interactions by measuring the released or absorbed heat during the process. The binding constant K can be determined directly from a single experiment using a non-linear least squares fit of the calorimetric titration data. Nevertheless, it is not possible to determine the constants of metal ion binding to apo-MTs as the heat change associated with this process contains both contributions from metal–ligand interaction as well as from protein folding, the latter being inherently linked to metal ion coordination in these otherwise mainly unstructured proteins. Accordingly, these two concomitant processes cannot be separated. Therefore, only the calorimetric titration of Zn_4MT3 with Cd^{II} was performed assuming isostructural replacement of Zn^{II} ions by Cd^{II} and hence a negligible contribution of protein folding/rearrangement to the observed heat change (SI). To avoid any kind of oxidation, the experiment was carried out inside an anaerobic chamber and all solutions were carefully degassed. The data can be fitted with a model considering a single binding site, probably because Cd^{II} is inserted into the preformed Zn_4 cluster resulting in relatively similar K values for all binding sites, i.e. $(3.06 \pm 0.16) \times 10^5\ M^{-1}$. This value is roughly in the range of the constant for Zn^{II}/Cd^{II} exchange observed for equine liver MT ($10^4\ M^{-1}$) [45]. The slightly higher value might reflect the apparent higher stability of Cd_4musMT3 versus Zn_4musMT3 that goes beyond the “normal” increase always observed in MTs due to the intrinsic higher thermodynamic stability of Cd^{II}-thiolate bonds compared to Zn^{II}-thiolate bonds: when both species are produced in Cd^{II} or Zn^{II}-supplemented *E. coli* media, a Cd_4- but only a Zn_3-species is isolated suggesting a reduced stability of the fourth Zn^{II} ion [13].

Conclusions

The presented data provide strong evidence that the metalation pathway of musMT3 with divalent d^{10} metal ions proceeds via a M_3^{II}-species, in which the metal ions are coordinated separately to the N- and C-terminal Cys-rich regions. The fourth metal ion leads to a decrease of the hydrodynamic radius suggesting a globular arrangement with connected N- and C-terminal regions. To our knowledge, the observation that a MT with two spatially separated Cys-rich regions can form two discrete species that show distinct differences both in metal ion load and in hydrodynamic radii is unprecedented so far. In addition, the UV/vis data from the Cd^{II} and Co^{II} titrations suggest that coordination of the first two metal ions occurs without recruitment of bridging thiolate ligands, which is only feasible assuming binding of one metal ion each to the N- and C-terminal Cys-rich regions (Fig. 11a).

Concerning the role of the His residue, the EXAFS data cannot by themselves exclude the possibility of exclusively thiolate ligation, but they are consistent with the coordination of seven Cys and a single His to the two Zn^{II} ions, as suggested by the visible absorption spectra. In addition, the EXAFS data provide reasonable metrical parameters for a purported ZnS_3N environment that may be used to model the average Zn^{II} environment in the truncated HC_6 protein.

Analyses based on the Co^{II} spectra of the separate domains in the visible region reveal as a most likely scenario the occurrence of mixed species with and without His as a ligand with a ratio of roughly 1:4.0 (Fig. 11b). Figure 11c, d depict possible cluster structures for the Zn_4-species for both arrangements either considering a minimal rearrangements of ligands compared to the proposed structure of the M_3^{II} species (top) or considering a large similarity to reported structures in the literature (bottom) [46, 47]. Structure predictions with I-TASSER show H46 to be involved in a hydrogen bond with E42 and might indicate a structural role. Unfortunately, I-TASSER does not address metal ion coordination, so an alternative possibility is that this His residue could coordinate at a reasonable distance to the divalent metal ion.

Generally, while examples for His-containing MT sequences have become more frequent in recent years, the role of this potential metal ion binding residue is still poorly understood. On the one hand, His has been identified as a ligand for Zn^{II} ions in a bacterial and a plant MT [11, 12]. On the other hand, non-coordinating His residues have also been described, which, nevertheless, might be important for structure stabilization: a non-coordinating His residue in the aforementioned bacterial MT can be envisioned, with minimal rearrangement, to be involved in a hydrogen bond to the backbone carbonyl oxygen of a neighbouring loop (His40-Cys47) [11], and the N-terminal His residue in the CupI MT from baker's yeast is engaged in a structure stabilizing hydrogen bond of 2.7 Å length in the solid state structure [48]. If the dual arrangement of the rather conserved His residue in musMT3 proposed here has any functional relevance for musMT3 or if it is "just" resulting from a high degree of conformational flexibility in this protein (which might be in itself functionally important) is a very intriguing but still hypothetical and completely unresolved question.

In particular, in light of the results obtained for the C-terminal HC_6H mutant that seem to support a persisting metal–His coordination, a further investigation of this peptide as well as of the corresponding full-length mutant seems highly interesting and will be the subject of further investigations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors acknowledge funding from the Swiss National Science Foundation (E.F.) and the National Institutes of Health (V.L.P., ES012236; A.D. and J.E.P.-H., GM 38047) for support of this research. Synchrotron measurements were made at the Stanford Synchrotron Radiation Laboratory, which is supported by the NIH Research Resource Program and the US Department of Energy.

References

1. Binz, P-A., Kägi, JHR. Metallothionein iv. Klaassen, C., editor. Birkhäuser; Verlag, Basel: 1999. p. 7-13.
2. Margoshes M, Vallee BL. J Am Chem Soc. 1957; 79:4813–4814.
3. Vallee, BL. Metallothionein. Kägi, JHR., Nordberg, M., editors. Birkhäuser; Verlag, Basel: 1979. p. 19-40.
4. Palmiter RD. Proc Natl Acad Sci USA. 1998; 95:8428–8430. [PubMed: 9671693]
5. Kojima, Y., Binz, P-A., Kägi, JHR. Metallothionein iv. Klaassen, C., editor. Birkhäuser; Verlag, Basel: 1999. p. 3-6.
6. Freisinger E. Dalton Trans. 2008; 17:6663–6675.
7. Freisinger E. Met Ions Life Sci. 2009; 5:107–153.
8. Freisinger E. Chimia. 2010; 64:217–224. [PubMed: 21138186]
9. Freisinger E. J Biol Inorg Chem. 2011; 16:1035–1045. [PubMed: 21688177]
10. Cobbett C, Goldsbrough P. Annu Rev Plant Biol. 2002; 53:159–182. [PubMed: 12221971]
11. Blindauer CA, Harrison MD, Parkinson JA, Robinson AK, Cavet JS, Robinson NJ, Sadler PJ. Proc Natl Acad Sci USA. 2001; 98:9593–9598. [PubMed: 11493688]
12. Peroza EA, Schmucki R, Güntert P, Freisinger E, Zerbe O. J Mol Biol. 2009; 387:207–218. [PubMed: 19361445]
13. Freisinger E. Inorg Chim Acta. 2007; 360:369–380.
14. http://www.uniprot.org/uniprot/
15. Legge GB, Martinez-Yamout MA, Hambly DM, Trinh T, Lee BM, Dyson HJ, Wright PE. J Mol Biol. 2004; 343:1081–1093. [PubMed: 15476823]
16. Peroza EA, Al Kaabi A, Meyer-Klaucke W, Wellenreuther G, Freisinger E. J Inorg Biochem. 2009; 103:342–353. [PubMed: 19111340]
17. Dage JL, Sun HJ, Halsall HB. Anal Biochem. 1998; 257:176–185. [PubMed: 9514787]
18. Hnizda A, Santrucek J, Sanda M, Strohalm M, Kodicek M. J Biochem Biophys Method. 2008; 70:1091–1097.
19. Miles EW. Methods Enzymol. 1977; 47:431–442. [PubMed: 22021]
20. Tomas M, Pagani MA, Andreo CS, Capdevila M, Bofill R, Atrian S. J Biol Inorg Chem. 2014; 19:1149–1164. [PubMed: 24951240]
21. Good M, Vašák M. Biochemistry. 1986; 25:3328–3334. [PubMed: 3524678]
22. Vašák M, Kägi JHR, Holmquist B, Vallee BL. Biochemistry. 1981; 20:6659–6664. [PubMed: 6272845]
23. Krizek BA, Merkle DL, Berg JM. Inorg Chem. 1993; 32:937–940.
24. Guo JQ, Giedroc DP. Biochemistry. 1997; 36:730–742. [PubMed: 9020770]
25. Kapust RB, Tözser J, Fox JD, Anderson DE, Cherry S, Copeland TD, Waugh DS. Protein Eng. 2001; 14:993–1000. [PubMed: 11809930]
26. Huber T, Freisinger E. Dalton Trans. 2013; 42:8878–8889. [PubMed: 23636452]
27. Pedersen AO, Jacobsen J. Eur J Biochem. 1980; 106:291–295. [PubMed: 7341229]
28. Wan X, Freisinger E. Metallomics. 2009; 1:489–500. [PubMed: 21305157]
29. Jelesarov I, Bosshard HR. J Mol Recognit. 1999; 12:3–18. [PubMed: 10398392]
30. Wiseman T, Williston S, Brandts JF, Lin LN. Anal Biochem. 1989; 179:131–137. [PubMed: 2757186]
31. Peroza EA, Freisinger E. Protein Expres Purif. 2008; 57:217–225.
32. Freisinger E, Vašák M. Met Ions Life Sci. 2013; 11:339–371. [PubMed: 23430778]
33. Irvine GW, Pinter TBJ, Stillman MJ. Metallomics. 2016; 8:71–81. [PubMed: 26583802]
34. Duncan KER, Stillman MJ. J Inorg Biochem. 2006; 100:2101–2107. [PubMed: 17055583]
35. Kowald GR, Sturzenbaum SR, Blindauer CA. Int J Mol Sci. 2016; 17:65.
36. Peroza EA, Freisinger E. J Biol Inorg Chem. 2007; 12:377–391. [PubMed: 17211631]
37. Zhang Y. BMC Bioinform. 2008; 9:40.

38. http://www.rcsb.org/pdb/
39. Kille P, Winge DR, Harwood JL, Kay J. FEBS Lett. 1991; 295:171–175. [PubMed: 1765150]
40. Vašák M, Kägi JHR. Proc Natl Acad Sci USA. 1981; 78:6709–6713. [PubMed: 6273885]
41. Bertini I, Luchinat C, Messori L, Vašák M. J Am Chem Soc. 1989; 111:7296–7300.
42. Overnell J, Good M, Vašák M. Eur J Biochem. 1988; 172:171–177. [PubMed: 2831057]
43. Loebus J, Peroza EA, Blüthgen N, Fox T, Meyer-Klaucke W, Zerbe O, Freisinger E. J Biol Inorg Chem. 2011; 16:683–694. [PubMed: 21437709]
44. Clark-Baldwin K, Tierney DL, Govindaswamy N, Gruff ES, Kim C, Berg J, Koch SA, Penner-Hahn JE. J Am Chem Soc. 1998; 120:8401–8409.
45. Vašák M, Kägi JHR. Met Ions Biol Syst. 1983; 15:213–273.
46. Braun W, Vašák M, Robbins AH, Stout CD, Wagner G, Kägi JHR, Wüthrich K. Proc Natl Acad Sci USA. 1992; 89:10124–10128. [PubMed: 1438200]
47. Henkel G, Krebs B. Chem Rev. 2004; 104:801–824. [PubMed: 14871142]
48. Calderone V, Dolderer B, Hartmann HJ, Echner H, Luchinat C, Del Bianco C, Mangani S, Weser U. Proc Natl Acad Sci USA. 2005; 102:51–56. [PubMed: 15613489]

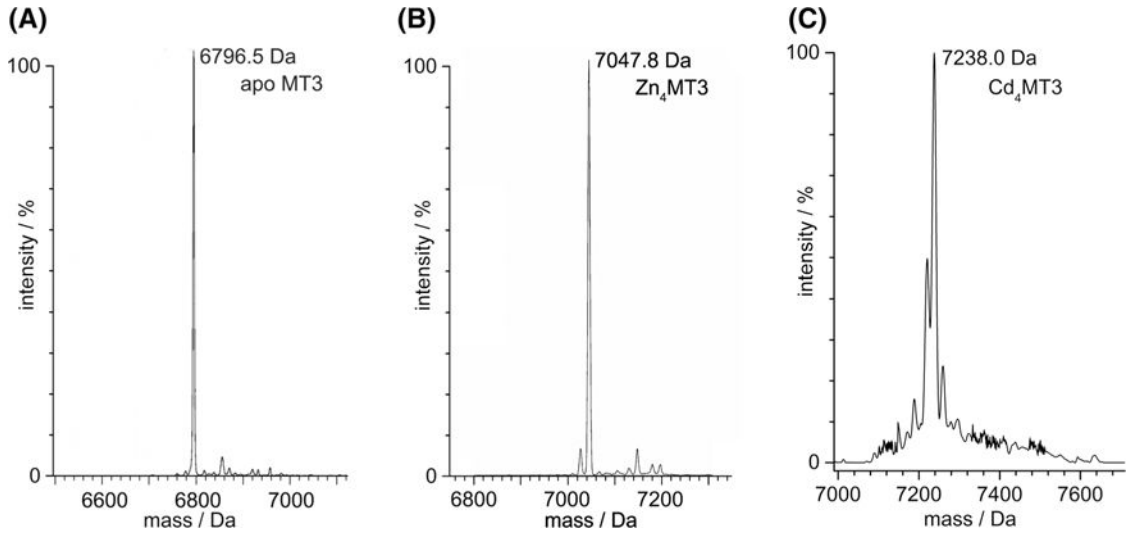

Fig. 1.

Deconvoluted ESI–MS spectra of **a** apo-musMT3 (calc. 6796.6 Da) as well as **b** Zn_4-(calc. 7048.1 Da) and **c** Cd_4musMT3 (calc.7236.1 Da)

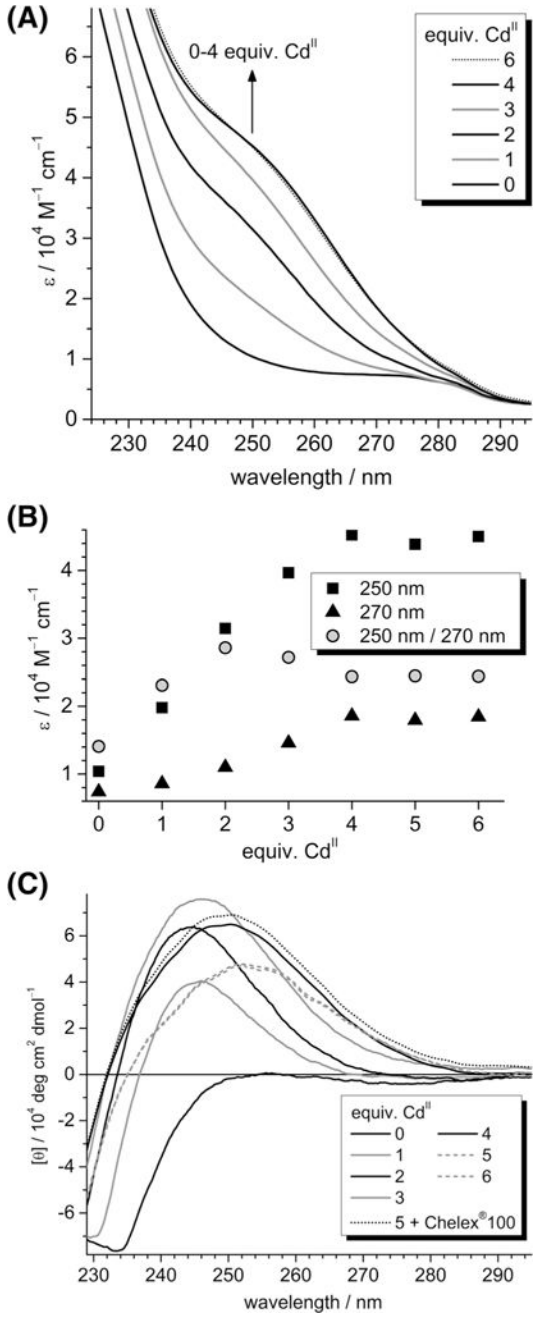

Fig. 2.

a UV spectra of the titration of apo-musMT3 with Cd^{II} ions in 1 mM Tris–HCl buffer (pH 7.5) and 10 mM NaCl. **b** Plot of extinction values at 250 and 270 nm as well as of the quotient of both values against the equiv. of Cd^{II} ions added. See "Discussion" in the text. **c** Titration of apo-musMT3 with Cd^{II} ions in 1 mM Tris–HCl buffer (pH 7.5) and 10 mM NaCl followed by CD spectroscopy. Ellipticity changes are observed up to the addition of five equiv. of Cd^{II}. The fifth metal ion can be removed by incubation with Chelex® 100 resin restoring the spectra obtained with four equiv. of Cd^{II}

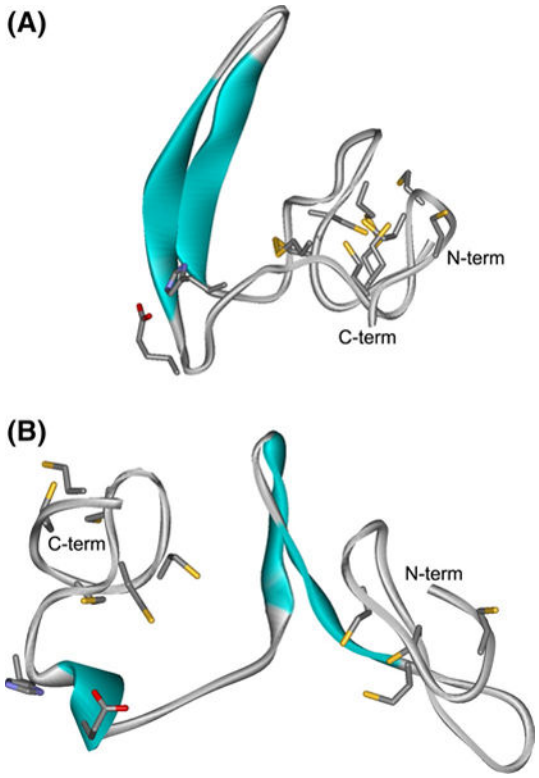

Fig. 3.

Models of the peptide fold for **a** Zn_4musMT3 and **b** Zn_3musMT3 as predicted by I-TASSER [37]. The peptide backbone is depicted as a grey ribbon, amino acid side chains are omitted for clarity except for the Cys residues (carbon and sulphur atoms only), H46 as well as E42. Secondary structural elements are depicted in cyan. For details of the prediction process see text

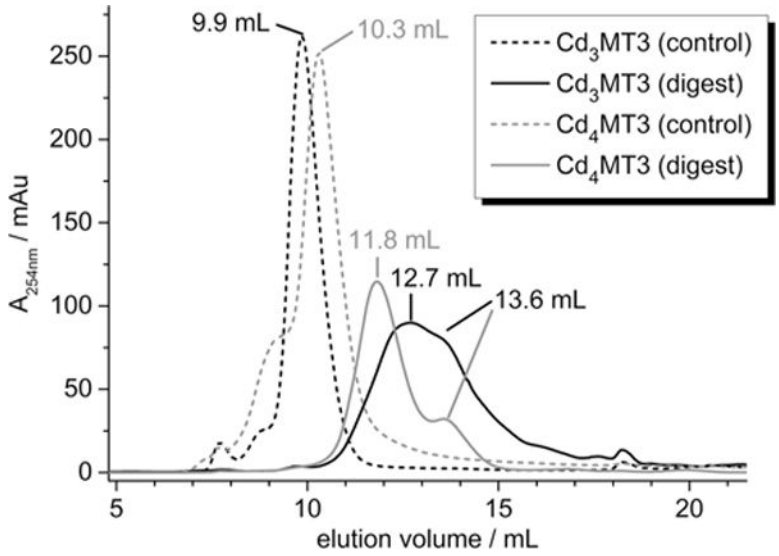

Fig. 4.

Gel filtration profile of undigested (dashed lines) and proteinase K digested (solid lines) Cd_3musMT3 (black lines) and Cd_4musMT3 (grey lines)

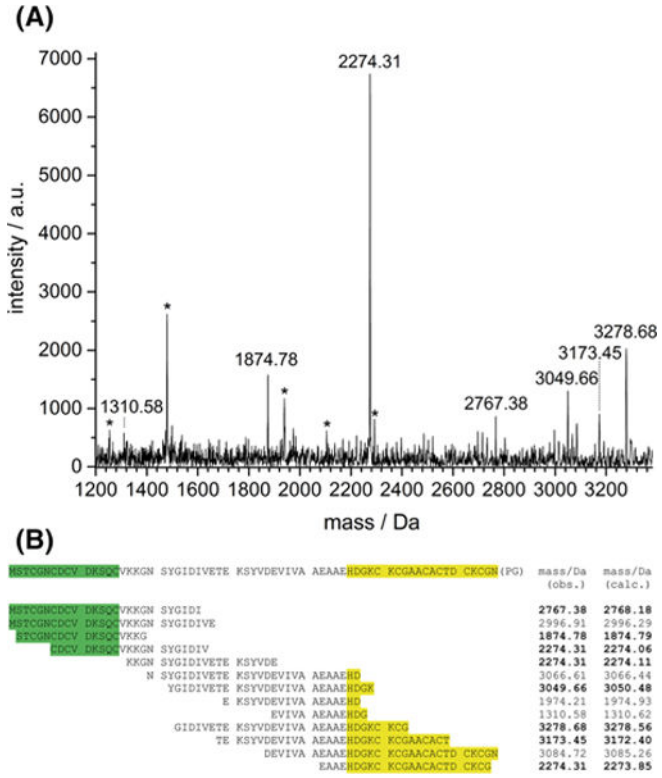

Fig. 5.

a MALDI-TOF of Cd_4musMT3 digested with *T. album* proteinase K. Signals marked with an asterisk (*) originate from proteinase K self-digestion. **b** Assignment of masses observed in (**a**) to peptide fragments of musMT3. Masses with high intensity observed in the spectra are marked in bold. Note that three different peptide fragments can be assigned to 2274.31 Da, which might contribute to the high intensity of the signal, keeping in mind that precise quantification of fragments is not possible from the mass spectra. The N-terminal Cys-rich region is highlighted in green, the C-terminal Cys-rich region in yellow

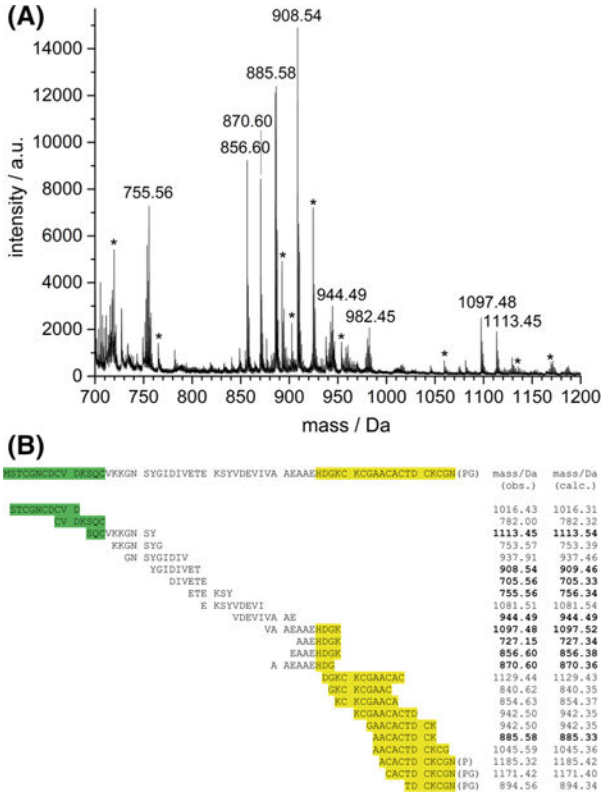

Fig. 6.

a MALDI-TOF of Cd_3musMT3 digested with *T. album* proteinase K. Signals marked with an asterisk (*) originate from proteinase K self-digestion. **b** Assignment of masses observed in (**a**) to peptide fragments of musMT3. Masses with high intensity observed in the spectra are marked in bold. The N-terminal Cys-rich region is highlighted in green, the C-terminal Cys-rich region in yellow

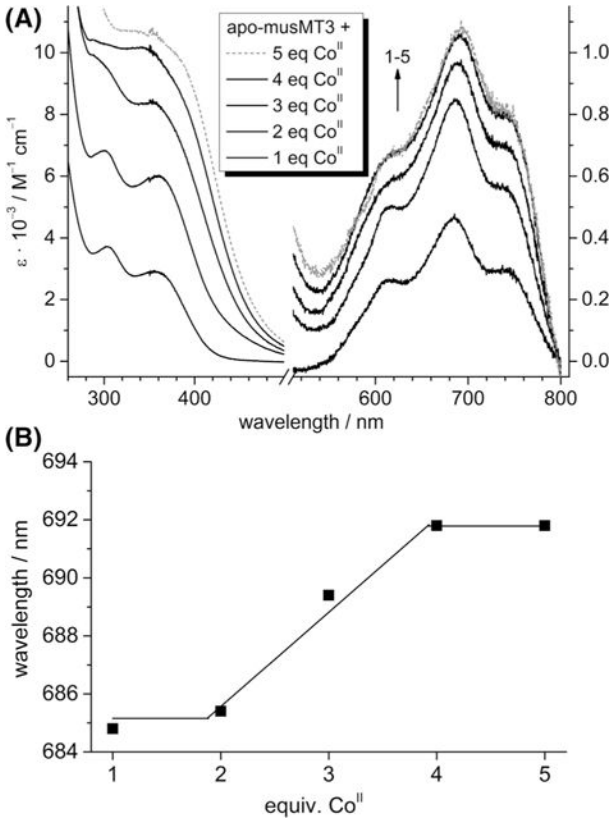

Fig. 7.

a UV/vis spectra of the titration of apo-musMT3 with Co^{II}. Note the different molar absorptivity scale for the UV (left) and the visible region (right). **b** Plot showing the bathochromic shift of the absorption band around 690 nm upon cluster formation between two and four equiv. of Co^{II}

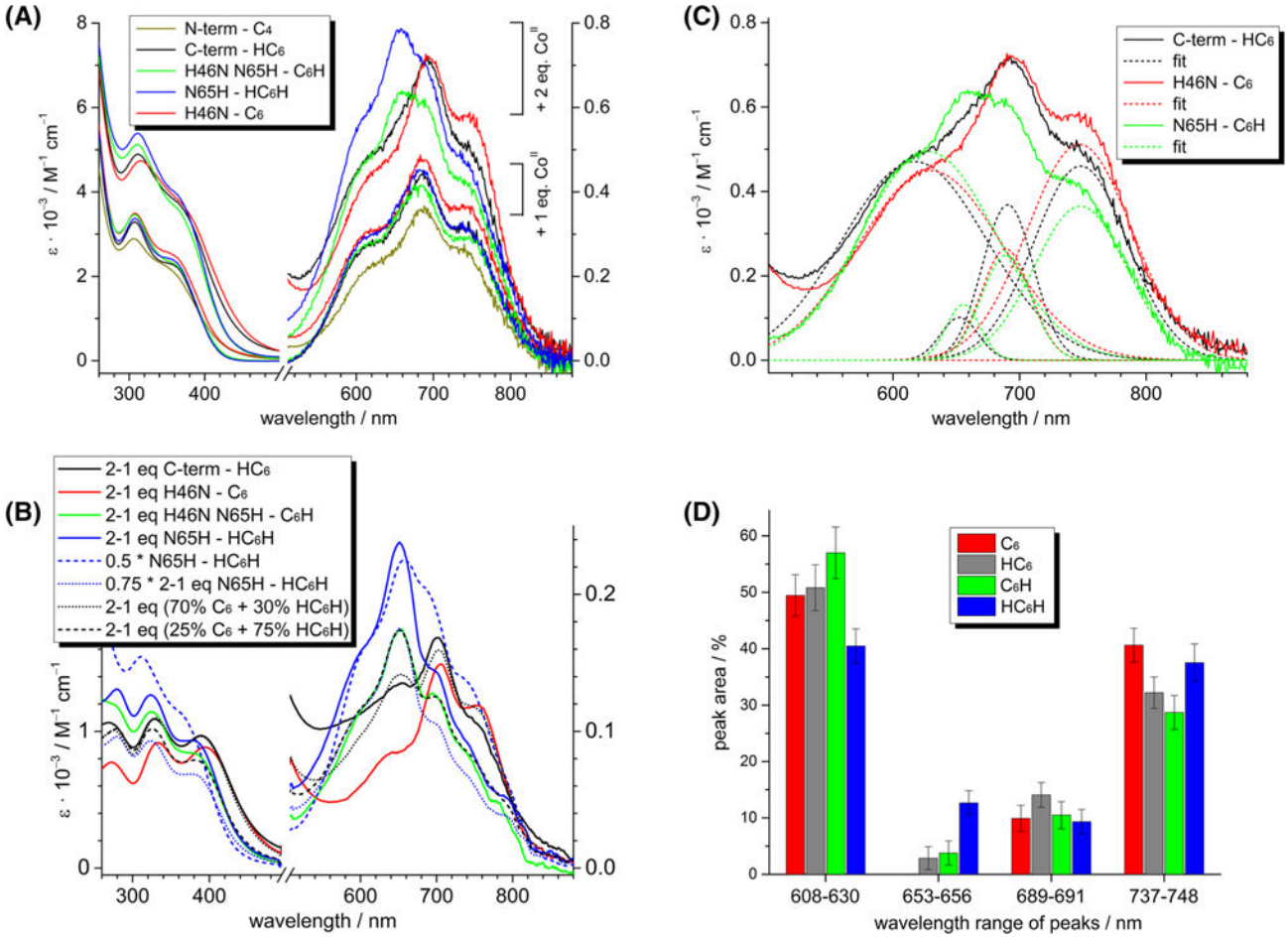

Fig. 8.

Co^{II} titration studies using the separate Cys-rich domains (C_4 and HC_6) as well as the three C-terminal mutants (C_6, C_6H, and HC_6H). **a** UV/vis spectra of the titration of the apo-peptides with one or two equiv. of Co^{II}. Note the different scale for the UV (left) and the visible region (right). **b** Difference spectra of the four C-terminal peptides obtained by subtracting the spectrum of the respective Co_1-species from the spectrum of the Co_2-species (solid lines). In addition, different approaches were applied to fit and explain the shape of the experimental spectra (dashed and dotted spectra). See text for "Discussion". **c** Fitting of the spectral envelope in the visible part of the spectra (solid lines) with three or four Gaussian peaks (dotted lines) to reveal the contribution of absorption bands to the overall spectra. **d** Column diagram of the Gaussian peak areas from (**c**) and SI versus the wavelength range of the respective peaks

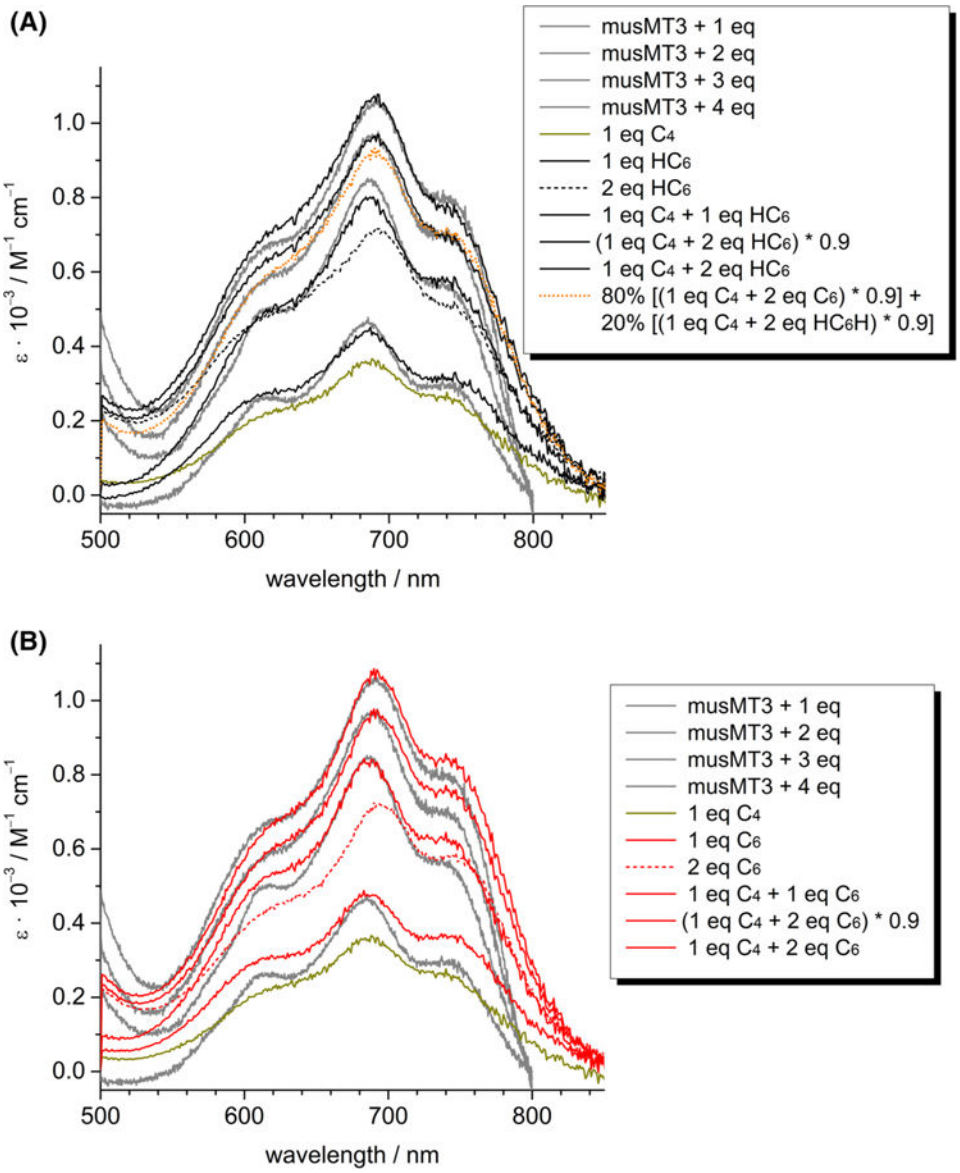

Fig. 9.

Fitting of the visible part of the $Co_{1\text{-}4}$musMT3 absorption spectra (grey solid lines) with the spectra of the separate Cys-rich regions. **a** Fittings with different combinations of spectra from the two wild-type sequences (C_4 and HC_6; olive and black spectra). Additionally, a fit using 80% of the C_6 mutant and 20% of the HC_6H mutant was performed (orange dotted line) and provides the best fit for the Co_3musMT3 data. **b** Fittings with different combinations of spectra from wild-type C_4 and the C-terminal C_6 mutant (olive and red lines). See text for detailed "Discussion"

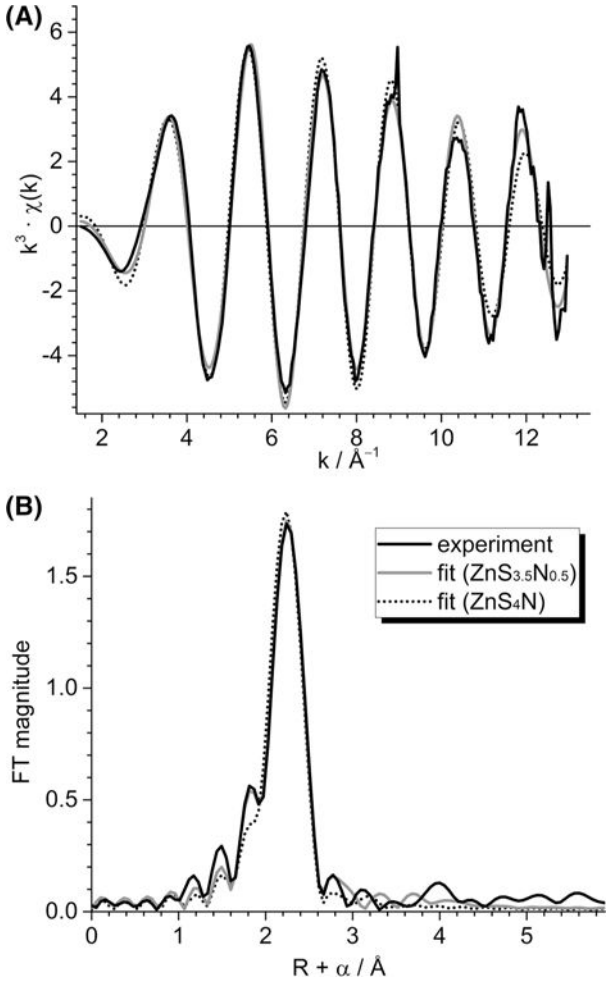

Fig. 10.

EXAFS spectroscopy of the wild-type C-terminal Cys-rich peptide Zn_2HC_6. **a** Experimental spectrum and fits over a range of $k = 1.5–13$ Å^{-1}. **b** Fourier transform of the EXAFS spectrum (over $k = 1.5–13$ Å^{-1}) including fits

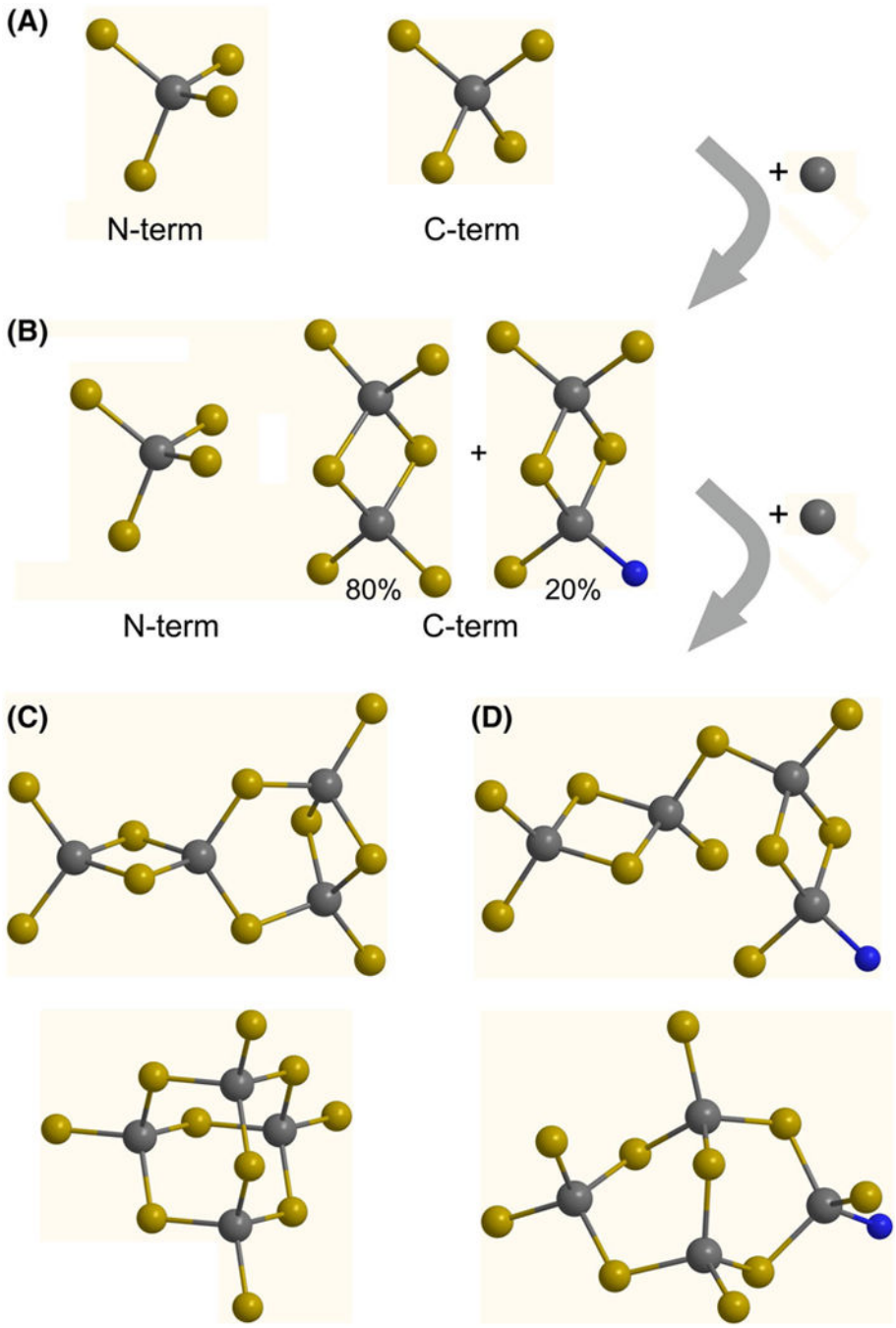

Fig. 11.

Proposed metalation pathway for Zn_4musMT3. **A** Coordination of the first two metal ions occurs separately in the N- and C-terminal Cys-rich regions. **b** The third metal ion is bound to the C-terminal part, recruiting one or more bridging thiolate ligands, and involves in roughly 20% of the molecules the His ligand. **c**, **d** Hypothetical cluster structures in Zn_4musMT3 without **(c)** and with consideration of the His ligand (**d**), requiring either

minimal rearrangements of ligands compared to (**b**) (top) or a rearrangement to yield structures similar to what is reported in the literature (bottom)

Table 1

Amino acid sequences of peptides used for the study of the separate Cys-rich domains

Name		Residues	Sequence
N-terminal Cys-rich domain			
WT	(C_4)	2–20	STCGNCDCV DKSQCVKKGN
C-terminal Cys-rich domain			
WT	(HC_6)	44–65	AEHDGKC KCGAACACTD CKCGN
H46N	(C_6)	44–65	AENDGKC KCGAACACTD CKCGN
H46N N65H	(C_6H)	44–65	AENDGKC KCGAACACTD CKCGH
N65H	(HC_6H)	44–65	AEHDGKC KCGAACACTD CKCGH